Split Metrics: Theory and Applications

M. Zahri

Department of Mathematics
Taibah University, Madinah

Seminars of the Department of Mathematics
1434/2013

Outline

1. Motivation
   - Sturmfels Inquiry
   - Bandelt and Dress Metric decomposition
2. Split metric
   - Definition of split pseudometric
   - Example of split or cut metric
3. Construction of the metric cone
   - Metric vector an card.
   - Metric cone
4. Extremal pseudometric
   - Extremal Rays
   - Split metric characterization in n-points space
   - From Extremal metrics to DNA property
5. Simulation of a class of random metrics
   - Random metric matrix and vector
   - Simulation of Random Metrics

Outline

1. Motivation
   - Sturmfels Inquiry
   - Bandelt and Dress Metric decomposition
2. Split metric
   - Definition of split pseudometric
   - Example of split or cut metric
3. Construction of the metric cone
   - Metric vector an card.
   - Metric cone
4. Extremal pseudometric
   - Extremal Rays
   - Split metric characterization in n-points space
   - From Extremal metrics to DNA property
5. Simulation of a class of random metrics
   - Random metric matrix and vector
   - Simulation of Random Metrics

Biology and Mathematics

- Can Biology Lead To New Theorems? Bernd Sturmfels in [5]. DNA and Evolution Phenomena (Dress Theorem for Example)
- Mathematics is biology's next microscope, only better; biology is mathematics' next physics, only better. J.E. Cohen in [2]
- Where are the new theorems? Bernd Sturmfels asks in [5]!

Outline

1. Motivation
   - Sturmfels Inquiry
   - Bandelt and Dress Metric decomposition
2. Split metric
   - Definition of split pseudometric
   - Example of split or cut metric
3. Construction of the metric cone
   - Metric vector an card.
   - Metric cone
4. Extremal pseudometric
   - Extremal Rays
   - Split metric characterization in n-points space
   - From Extremal metrics to DNA property
5. Simulation of a class of random metrics
   - Random metric matrix and vector
   - Simulation of Random Metrics

Decomposition of Metric on FM

Bandelt and Dress: A canonical decomposition theory for metrics on a finite set, Advances in Mathematics 92 (1992):

Theorem (Bandelt& Dress)

Every finite metric space D admits a unique coherent decomposition

$$D = D' + \sum_{i=1}^{k} D_i$$

where D_i, $i = 1, \ldots, k$ are linearly independent split metrics and D' is a split-prime metric.

- Split or Cut metric admits two values only! (next slide)
- Split-Prim metric are undecomposable metrics in term of BD-Theorem.

Outline

1. Motivation
   - Sturmfels Inquiry
   - Bandelt and Dress Metric decomposition
2. Split metric
   - Definition of split pseudometric
   - Example of split or cut metric
3. Construction of the metric cone
   - Metric vector an card.
   - Metric cone
4. Extremal pseudometric
   - Extremal Rays
   - Split metric characterization in n-points space
   - From Extremal metrics to DNA property
5. Simulation of a class of random metrics
   - Random metric matrix and vector
   - Simulation of Random Metrics

Split metrics, definition

Definition

Let A be a subset of a finite set X. The split (or cut) metric is a mapping defined as

$$d(x,y) := \begin{cases} 0, & \text{if } (x,y) \in A \times A \text{ or } A^c \times A^c \\ 1, & \text{otherwise.} \end{cases}$$

Where A^c denote the complement set of A in X.

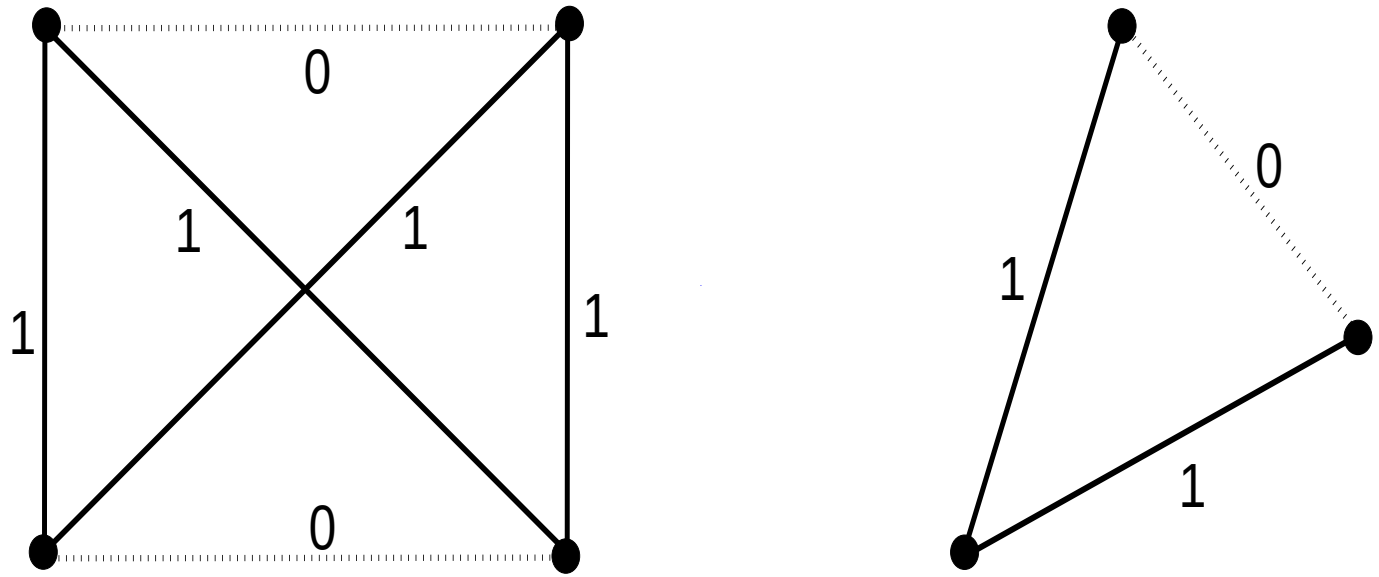
Note that a pseudometric space is a generalization of a metric space in which we allow the possibility that $d(x,y) = 0$ for distinct values of x and y. Thus a pseudometric is a metric if $d(x,y) > 0$ for $x \neq y$.

Outline

1. Motivation
   - Sturmfels Inquiry
   - Bandelt and Dress Metric decomposition
2. Split metric
   - Definition of split pseudometric
   - Example of split or cut metric
3. Construction of the metric cone
   - Metric vector an card.
   - Metric cone
4. Extremal pseudometric
   - Extremal Rays
   - Split metric characterization in n-points space
   - From Extremal metrics to DNA property
5. Simulation of a class of random metrics
   - Random metric matrix and vector
   - Simulation of Random Metrics

Example of Split Metrics.

Consider the finite point space: $X = \{1, 2, 3\}$ or $X = \{1, 2, 3, 4\}$



Example of splitting of three and four points space

Outline

1. Motivation
   - Sturmfels Inquiry
   - Bandelt and Dress Metric decomposition
2. Split metric
   - Definition of split pseudometric
   - Example of split or cut metric
3. Construction of the metric cone
   - Metric vector an card.
   - Metric cone
4. Extremal pseudometric
   - Extremal Rays
   - Split metric characterization in n-points space
   - From Extremal metrics to DNA property
5. Simulation of a class of random metrics
   - Random metric matrix and vector
   - Simulation of Random Metrics

Metric Vector & Card.

Definition

We call d_v a pseudometric **vector**, which is the C_n^2 dimensional vector set of vectors of the form

$$d_v = \left(d_{1,2}, \ldots, d_{1,n}, d_{2,3}, \ldots, d_{2,n}, \ldots, d_{n-1,n}, d_{n,n}\right),$$

where $d_{i,j}$ are the entries of the corresponding pseudometric d.

Definition

The triple $(d_{i,j}, d_{j,k}, d_{i,k}))$, for all $i, j, k = 1 \ldots, m$, is called pseudometric **card** and the triple (x_i, x_j, x_k) is called an admissible card if and only if $x_i, x_j, x_k \geq 0$ and $x_i \leq x_j + x_k$; $\quad \forall i, j, k = 1, \ldots, m$.

Outline

1. Motivation
   - Sturmfels Inquiry
   - Bandelt and Dress Metric decomposition
2. Split metric
   - Definition of split pseudometric
   - Example of split or cut metric
3. Construction of the metric cone
   - Metric vector an card.
   - Metric cone
4. Extremal pseudometric
   - Extremal Rays
   - Split metric characterization in n-points space
   - From Extremal metrics to DNA property
5. Simulation of a class of random metrics
   - Random metric matrix and vector
   - Simulation of Random Metrics

Metric Cone

Consider the sets:
$R = \{(i,j,k) | i,j,k = 1,\ldots,n \text{ pairwise distinct}\}, \quad I = \{t | t = 1,\ldots,m\}.$

- Where R represents the indices of all admissible cards, which satisfies the metric axioms, and
- I refers the set of indices, such that the t^{th} components x_t represent a distance between two different points.

Definition

For $n \in \mathbb{N}$, $m = C_n^2$, $t = 1,\ldots,m$ and $i,j,k \in \{1,\ldots,n\}$, the maps l_t and $l_{i,j,k}$ defined as: $l_t,\ l_{i,j,k} : \mathbb{R}^m \longrightarrow \mathbb{R}$

$$l_t(x) = l_i(x_1,\ldots,x_m) = x_i, \quad l_{i,j,k}(x_1,\ldots,x_m) = x_j + x_k - x_i.$$

are linear, for $t \in I$ and $(i,j,k) \in R$.

Hence, we define the halfspaces K_t and $K_{i,j,k}$ as follows

$$K_t = \{x \in \mathbb{R}^m | \ \ell_t(x) \geq 0\}, \quad (1)$$

$$K_{i,j,k} = \{x \in \mathbb{R}^m | \ \ell_{i,j,k}(x) \geq 0\}. \quad (2)$$

Lemma

The following subsets

$$H_t = \{x \in \mathbb{R}^m | l_t(x) = 0\} \quad H_{i,j,k} = \{x \in \mathbb{R}^m | l_{i,j,k}(x) = 0\}, \quad (3)$$

are hyperplane in $\mathbb{R}^m$, defined by the extremal conditions of H_t and $H_{i,j,k}$.

We call the hyperplanes intersection above Metric cone.

Outline

1. Motivation
   - Sturmfels Inquiry
   - Bandelt and Dress Metric decomposition
2. Split metric
   - Definition of split pseudometric
   - Example of split or cut metric
3. Construction of the metric cone
   - Metric vector an card.
   - Metric cone
4. Extremal pseudometric
   - Extremal Rays
   - Split metric characterization in n-points space
   - From Extremal metrics to DNA property
5. Simulation of a class of random metrics
   - Random metric matrix and vector
   - Simulation of Random Metrics

Extremal pseudometric

Definition

A pseudometric d is called extremal metric if for all $g, h \in M_0$, the following holds: $g + h = d$ implies $g = \nu d$, $h = \mu d$ with $\nu, \mu \geq 0$. If d is extremal, then $(\alpha d)_{\alpha \in \mathbb{R}^+}$ is called extremal ray of d.

It is important to note that the notion of extremal metric is equivalent to the notion of extremal metric vector. Therefore, we deal with metric vectors.

Lemma

An extremal pseudometric vector is given by the intersection of $m - 1 = C_m^2 - 1$ *of the hyperplanes* H_t *and* $H_{i,j,k}$.

Outline

1. Motivation
   - Sturmfels Inquiry
   - Bandelt and Dress Metric decomposition
2. Split metric
   - Definition of split pseudometric
   - Example of split or cut metric
3. Construction of the metric cone
   - Metric vector an card.
   - Metric cone
4. Extremal pseudometric
   - Extremal Rays
   - Split metric characterization in n-points space
   - From Extremal metrics to DNA property
5. Simulation of a class of random metrics
   - Random metric matrix and vector
   - Simulation of Random Metrics

Theorem

Lemma (Zahri 2011)

Let (X, d) be a n points metric space and if EX_n is the set of all extremal pseudometric vectors, then $n\Big(\{d|d \in EX_n\}\Big) = 2^{n-1}$.

Theorem (Zahri 2011)

Every split metric of a n points metric space is extremal and if $n < 5$, then the unique extremal pseudometrics are the split metrics.

- To the proof: We check that the split metrics satisfies the required number of Hyperplanes intersection.
- If $n \leq 4$, we have $d = \sum_{i=1}^{2^{n-1}} \alpha_i d_i$, where d_i is a split metric and $\alpha_i \in \mathbb{R}^+$

Split metric vectors for n=3,4 and 5

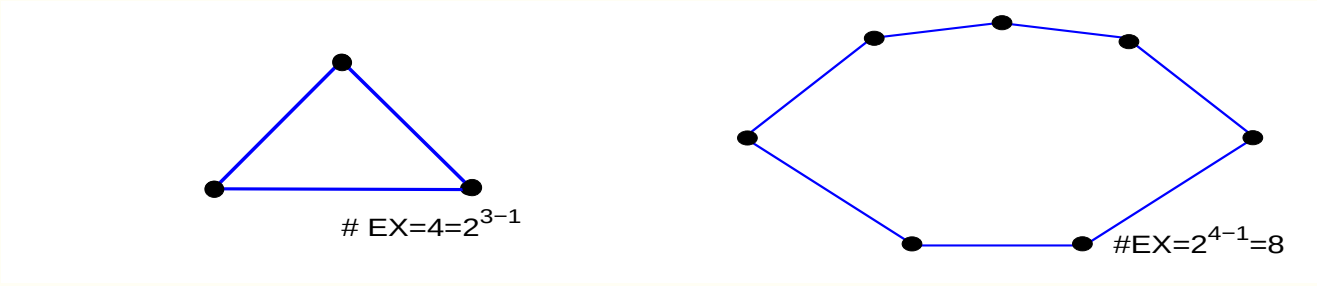
$$EX_3 = \left\{(0,0,0);(0,1,1);(1,0,1);(1,1,0)\right\}$$

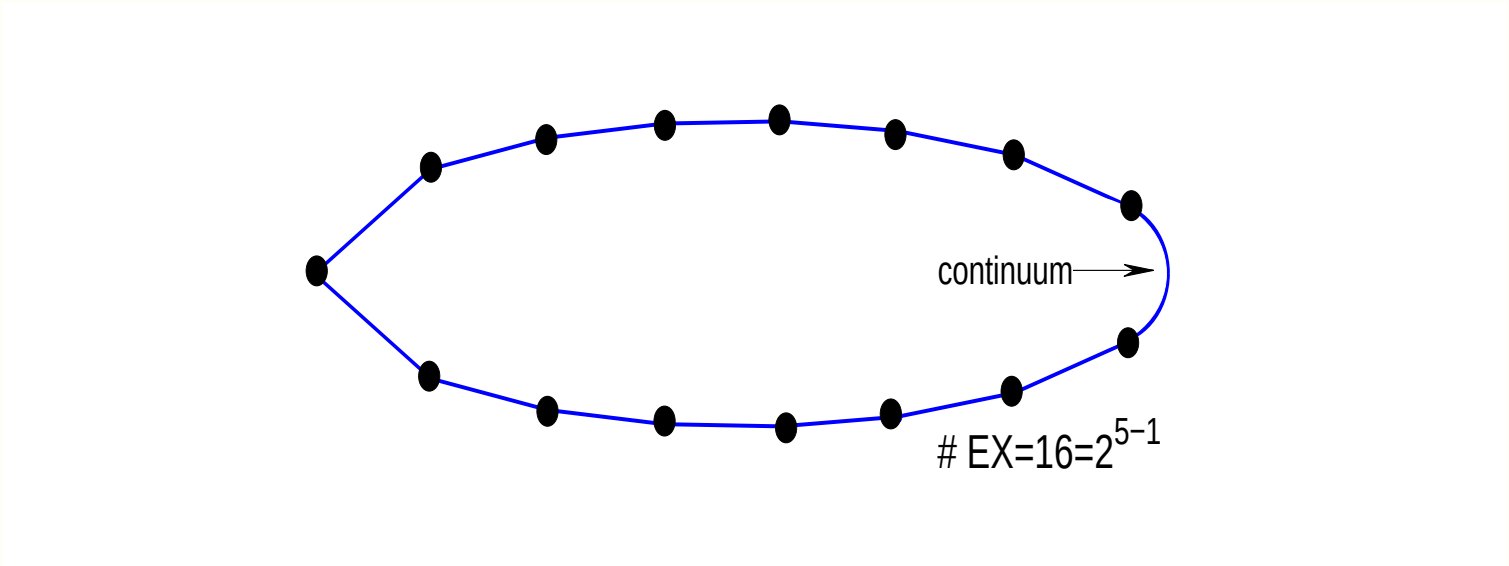
$$EX_4 = \left\{(0,0,0,0,0,0);(0,0,1,0,1,1);(0,1,0,1,0,1);(1,0,0,1,1,0);\right.$$
$$\left.(1,1,1,0,0,0);(0,1,1,1,1,0);(1,0,1,1,0,1);(1,1,0,0,1,1)\right\}$$

$$EX_5 = \left\{(0,\ldots,0),\underbrace{(1,1,1,1,0,\ldots,0)}_{Splitting\ 1 from\ 2,3,4,5},\right.$$
$$\left.\ldots(d_{1,2},d_{1,3},d_{1,4},d_{1,5},\ldots,d_{4,5},)\right\} \quad (4)$$

Cut of the metric cone

Horizontal cut of the metric cones for n=3,4,5.





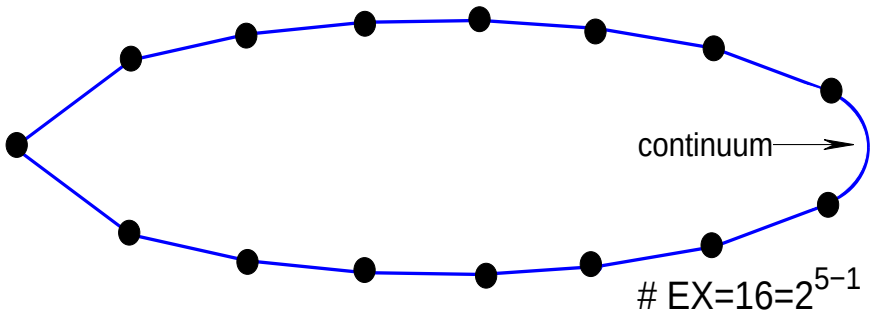
Remark that: ♯EX= zero point + extremal rays (dotes).

Outline

1. Motivation
   - Sturmfels Inquiry
   - Bandelt and Dress Metric decomposition
2. Split metric
   - Definition of split pseudometric
   - Example of split or cut metric
3. Construction of the metric cone
   - Metric vector an card.
   - Metric cone
4. Extremal pseudometric
   - Extremal Rays
   - Split metric characterization in n-points space
   - From Extremal metrics to DNA property
5. Simulation of a class of random metrics
   - Random metric matrix and vector
   - Simulation of Random Metrics

DNA property

- The undecomposable human DNA is a linear combination of the boundary of the extremal rays of the five point set.
- Because of the **continuum**, the DNA of two different humans are equal with probability zero. (Mathematical certainty)
- Both DNA molecules (Ribose, Deoxyribose) may be represented by the five edges (five points) graph.How could we explain this similarity!



Remark that: ♯EX= zero point + extremal rays (dotes).

Outline

1. Motivation
   - Sturmfels Inquiry
   - Bandelt and Dress Metric decomposition
2. Split metric
   - Definition of split pseudometric
   - Example of split or cut metric
3. Construction of the metric cone
   - Metric vector an card.
   - Metric cone
4. Extremal pseudometric
   - Extremal Rays
   - Split metric characterization in n-points space
   - From Extremal metrics to DNA property
5. Simulation of a class of random metrics
   - Random metric matrix and vector
   - Simulation of Random Metrics

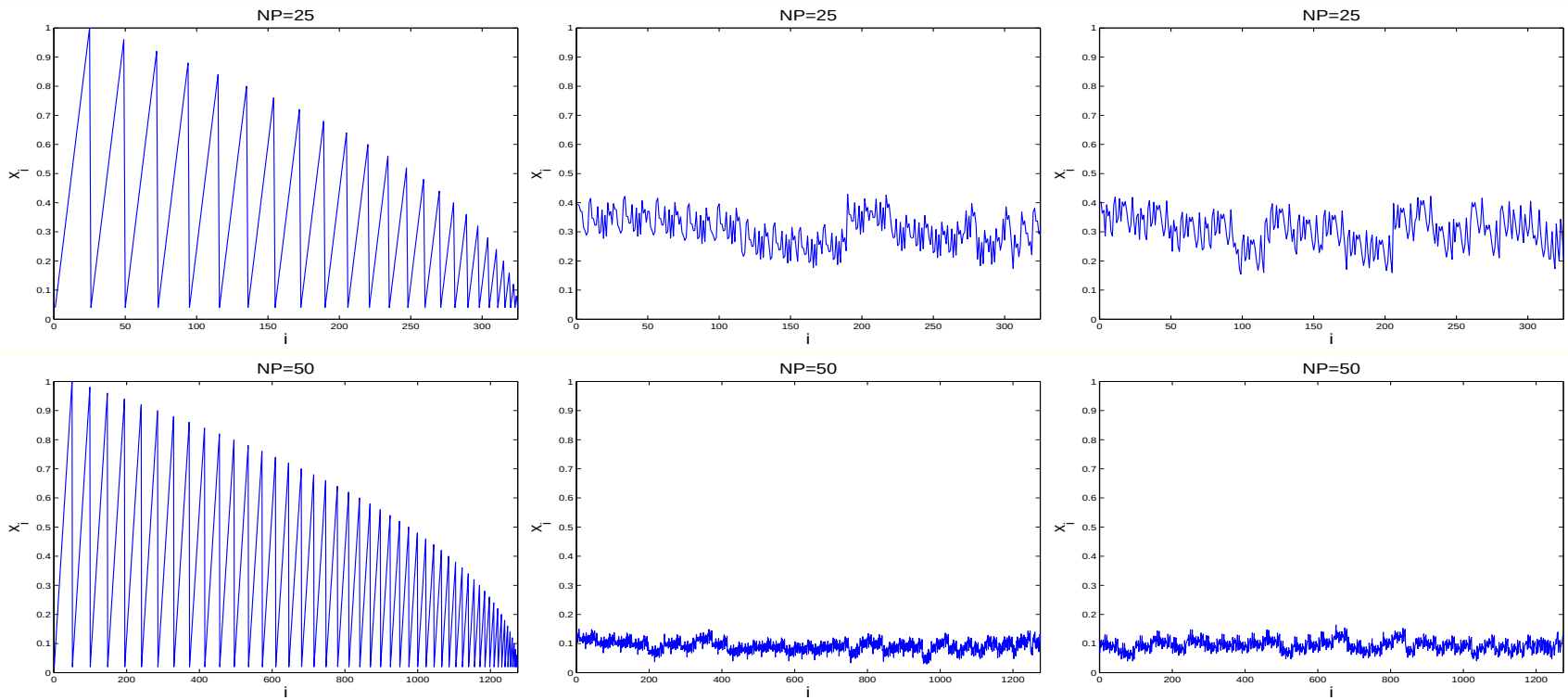
Random metric matrix

Generating a 8-point random metric matrix: $d_r =$

$$\begin{pmatrix}
0 & 0.5062 & 0.4627 & 0.5693 & 0.2864 & 0.4828 & 0.3735 & 0.3238 \\
0.5062 & 0 & 0.634 & 0.7406 & 0.4577 & 0.654 & 0.5447 & 0.4951 \\
0.4627 & 0.634 & 0 & 0.6971 & 0.4142 & 0.6106 & 0.5013 & 0.4516 \\
0.5693 & 0.7406 & 0.6971 & 0 & 0.5208 & 0.7172 & 0.6079 & 0.5582 \\
0.2864 & 0.4577 & 0.4142 & 0.5208 & 0 & 0.4343 & 0.325 & 0.2753 \\
0.4828 & 0.654 & 0.6106 & 0.7172 & 0.4343 & 0 & 0.5214 & 0.4717 \\
0.3735 & 0.5447 & 0.5013 & 0.6079 & 0.325 & 0.5214 & 0 & 0.3624 \\
0.3238 & 0.4951 & 0.4516 & 0.5582 & 0.2753 & 0.4717 & 0.3624 & 0
\end{pmatrix}$$

For the used algorithms see [6].

Random metric vector



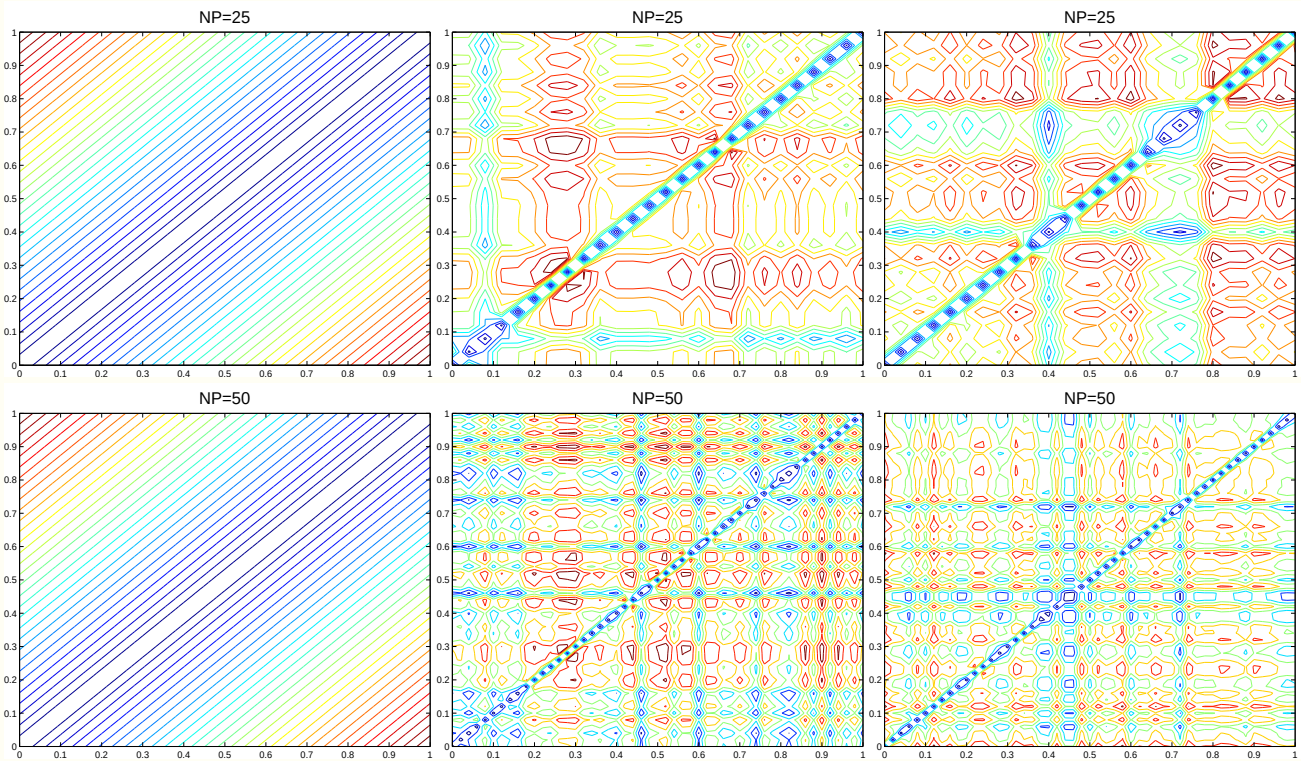
Euclidean metric vector (first column) and four simulation of random metric vectors for NP=25 and 50

Outline

1. Motivation
   - Sturmfels Inquiry
   - Bandelt and Dress Metric decomposition
2. Split metric
   - Definition of split pseudometric
   - Example of split or cut metric
3. Construction of the metric cone
   - Metric vector an card.
   - Metric cone
4. Extremal pseudometric
   - Extremal Rays
   - Split metric characterization in n-points space
   - From Extremal metrics to DNA property
5. Simulation of a class of random metrics
   - Random metric matrix and vector
   - Simulation of Random Metrics

Simulation of Random Metrics NP=25, 50

Something like **fingerprints**: DNA is like a fingerprint because everyone's is a little different!



Contour plot for the Euclidean metric matrix (first column) and four simulation of random metric matrices for NP=25 and 50. We show clearly the symmetry of the Metrics

Open Question& a remark from B. Sturmfels

- How could we characterize Split-Prim-Metrics!
- Given a Metric, determine the topology (nature) of the corresponding metric space!
- I think it would be interesting to see how your notion of randomness relates to the notion of genericity that underlies the CLASSIFICATION OF SIX-POINT METRICS [4]
  Best wishes, Bernd Sturmfels (to me)

Bibliography

- H.J. Bandelt and A. Dress, A canonical split decomposition theory for metrics on a finite set, AiM 92/47-105, 1992
- J.E. Cohen: Mathematics is biology′s next microscope, only better; biology is mathematics′ next physics, only better, PLOS Biology 2 (2004) No.12.
- A. Dress , Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: a note on combinatorial properties of metric spaces, Advances in Math. 53 no. 3, 321-402, 1984
- B. Sturmfels, Josephine Yu , Classification of Six-Point Metrics, The Electronic Journal of Combinatorics , Volume 11, Issue 1 (2004)
- B. Sturmfels, Can Biology Lead to New Theorems, posted at http://www.claymath.org/library/annual_report/Sturmfels.pdf
- M. Zahri, Construction and Simulation of a Class of Random Metrics on Finite Metric Spaces, J. of Num. Math.& Stoch., 4 (1) : 26-39, 2012